C400 0400



2

OIPE

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

Input Set: N:\Crf3\RULE60\10027806.raw
Output Set: N:\CRF3\01182002\J027806.raw

```
1 <110> APPLICANT: Swanson, Ronald V.
              Feldman, Robert A.
              Schleper, Christa
      4 <120> TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
      5 <130> FILE REFERENCE: DCORP.002A
      6 <140> CURRENT APPLICATION NUMBER: US/10/027,806
      7 <141> CURRENT FILING DATE: 2001-12-21
      8 <150> PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
W--> 9 <151> PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
     10 <160> NUMBER OF SEQ ID NOS: 123
     11 <170> SOFTWARE: FastSEO for Windows Version 3.0
     13 <210> SEO ID NO: 1
     14 <211> LENGTH: 32998
                                                          ENTERED
     15 <212> TYPE: DNA
     16 <213> ORGANISM: Cenarchaeum symbiosum
     17 <220> FEATURE:
     18 <221> NAME/KEY: CDS
     19 <222> LOCATION: (7604)...(8908)
     20 <221> NAME/KEY: CDS
     21 <222> LOCATION: (8961)...(9767)
     22 <221> NAME/KEY: CDS
     23 <222> LOCATION: (10545)...(10922)
     24 <221> NAME/KEY: CDS
     25 <222> LOCATION: (13944)...(14612)
     26 <221> NAME/KEY: CDS
     27 <222> LOCATION: (18638)...(20149)
     28 <221> NAME/KEY: CDS
     29 <222> LOCATION: (20554)...(20955)
     30 <221> NAME/KEY: CDS
     31 <222> LOCATION: (20956)...(21834)
     32 <221> NAME/KEY: CDS
     33 <222> LOCATION: (25151)...(26377)
     34 <221> NAME/KEY: CDS
     35 <222> LOCATION: (27535)...(28002)
     36 <221> NAME/KEY: CDS
     37 <222> LOCATION: (28065)...(29483)
     38 <400> SEQUENCE: 1
     39
              gateettgae etetgegett attgeageea tggaetgaee ggeegtgegg ggetaaataa
                                                                                      60
     40
              agetgaggeg eegeetgeag getetgetea geegttgatt atacagtaet egeactegea
                                                                                    120
     41
              ggtcttgctt tcgtcatctt ttgccctgca cttggcgtgc tggcccgact ggcactggac
                                                                                    180
     42
              gcagatctcg tctatgttgt ctatcgaggt gtactttttg gactggtcaa agccgaagcc
                                                                                     240
```

gccgtccttt ggtccaacca tgaagcggtc agggccgtta tgccttaatt atctttgccg

teggggeggg geogettetg eeggegggag eeceettga ggeegeteee eggetegttq

43

44

300

360

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

45			cctgatcata				420
46			cctagccacg				480
47			gcgcggggca				540
48			ggtaagctac				600
49			taatcatctg				660
50			ttgccagaaa				720
51			aggtacctac				780
52			gccagtcttg				840
53			tgcttttcac				900
54			tgtagttgag				960
55			tccttctctt				1020
56	caggacggtt	acatcatcgc	gtttgattat	ccgttctgtt	gtttcagctt	tttgctcatt	1080
57			acggtatttg				1140
58			taccggggga				1200
59	tgcatccacc	gtggacccat	gagacaacga	gccggcagcg	gtgcacagca	ccgagtacaa	1260
60			acgacctgcc				1320
61			tgagaattct				1380
62	caacgacggt	attaccgagg	tgattcatga	taattttgtt	gactggagta	tactgaaaca	1440
63			actgacagca				1500
64			atccttttac				1560
65			gacattattt				1620
66			tatctagcca				1680
67			tgaattagcc				1740
68			tgtttttctg				1800
69			atgtgcatgc				1860
70			tgtttaccca				1920
71			gcatatggga				1980
72			tcggtggact				2040
73	aatctaattt	cttagtcttt	aaatcacgca	tcatcctcat	aaatccatac	aggtcaccat	2100
74	ttgttatgag	agcttcatta	gacgtatgcg	cttcgtccgt	tatccagttg	gtcgcataca	2160
75			tactgttgca				2220
76			gactccatta				2280
77	aactaggcat	ttatgtagtt	gaaatgacta	cccgcgggaa	tcataccata	gtctgtgtcg	2340
78	tatgacttgc	cttttttctt	catcaatttc	tcatattcct	catgcagttc	gagcaggatg	2400
79	gccatgtcac	tacgtttgtt	cgtctgtttt	gttgtctcgg	gcttttggtc	catattatta	2460
80			tgttccttta				2520
81	ttgcctcgca	atttcctcca	aggcacatga	aaaacgggcc	acaggcagag	cacagcatcc	2580
82			gcccccggcg				2640
83	aaaatgagtc	atcatgcata	gtctctatgt	aaatggctga	accggtgttt	tggtcgatta	2700
84			cgccgaagct				2760
85	gtcttcttct	gcgaaccttc	atccgaagaa	ggaatatctt	gtctcgggat	aggattcgtg	2820
86			agcctagatg				2880
87			gcctctctgt				2940
88	gatattcgcg	cttccatcag	gcagaggccg	acctgtctca	cgacggtcta	aacccagctc	3000
89			agcagcctca				3060
90			ccaaaccgcg				3120
91			ttttctgtca				3180
92			gtctatgaat				3240
93	tttttggctt	tgccctcttc	agcggatttc	tgacccgctt	gaactaaact	ttgggcccct	3300

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

							2260
94	ttgatatctt						3360
95					tacatcggcg		3420
96	tcccaaagaa						3480
97					gatggaagat		3540
98					cagtggggct		3600
99					taccttaaga		3660
100					agttttagat		3720
101					gtcgcctgtg		3780
102					ttcctcatga		3840
103					ccctcgccat		3900
104						gaacttgcag	3960
105	tttgctagcc	gcacggtctt	tcatggtctc	ctggaatcgg	gaaaactctg	ctaacgcaaa	4020
106					cccaggccct		4080
107	gacacgacga	cggtcatgtt	ccccctatcc	ggaagcgaac	catgcggttc	aaacgctccc	4140
108	tgcaaggtac	cagaatatta	actggtttcc	cattcggact	actctgttga	ggcagtcctt	4200
109	aggatcgact	aactccaggc	tgacgacgca	ttgcctggaa	accettgege	ttacggtggt	4260
110	gcggattctc	accgcactat	gctgttactg	ccaccaggat	ctgcaataga	aatcggtcca	4320
111	caggacgtca	ccgccctgct	tcgtcccaat	cactacgcca	acctaccacg	gtgcacctat	4380
112	cacqqtqcac	gtctggagta	toggtactct	gctttagccc	cgtccgtttt	tgtggcgccc	4440
113						agcttacctc	4500
114					tagcagaaat		4560
115						gaacccgtgt	4620
116						atctcttccc	4680
117	-					ccctgcgaga	4740
118						attcccaagt	4800
119						cgcccacctt	4860
120						taacgcactt	4920
121						cgcctttcta	4980
122						aacgcatgac	5040
123						ccgaaaaaat	5100
124			_			acccccttc	5160
125						agagatattt	5220
126						ctactcgggt	5280
127		_	_			cggaacattt	5340
128						catctcggcc	5400
129						acttgggaaa	5460
130						gttcgacctc	5520
131	_						5580
						ggtgcgtgca	5640
132						ctagcaatcc	
133					tacacgacta		5700
134					tacaccactt		5760
135		-		_		ggtgatccga	5820
136						cctcaagttc	5880
137	-			-	. caatgaaacg		5940
138					gegeggttae		6000
139					ggtagcgttt		6060
140					agcccgcgtg		6120
141					ccgcattaac		6180
142	ccgctaattc	gccccactgc	tccggagagc	aatggtggca	. actagaggca	aggatctcgc	6240



RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/027,806

DATE: 01/18/2002
TIME: 09:17:38

143	togttacctg	acttaacagg	acatctcacg	gcacgagctg	gcgacggcca	tgcaccacct	6300
144	ctcagcttgt	ctggtagagt	cttcagcttg	accttcacac	tgctgtctct	ccgggtaaga	6360
145	tttctggcgt	tgactccaat	tgaaccgcag	gcttcacccc	ttgtggtgct	cccccgccaa	6420
146	ttcctttaaq	tatcatactt	gcgtacgtac	ttcccaggcg	gcaaacttaa	cggcttccct	6480
147	gcggcactgc	actggctctt	acgccaatgc	atcactgagt	ttgcattgtt	tacagctggg	6540
148	actacccggg	tatctaatcc	ggtttgctcc	cccagctttc	atccctcacc	gtcggacgtg	6600
149	ttctagtaga	ccgccttcgc	cacagggggt	catcgataga	tcagaggatt	ttaccccttc	6660
150	ctaccgagta	ccgtctacct	ctcccactcc	ctagccgtgc	agtatttccg	gcagcctatg	6720
151	cattgagcgc	atagatttaa	ccgaaaactt	acacggcagg	ctacggatgc	tttaggccca	6780
152	ataatcctcc	tgaccacttg	aggtgctggt	tttaccgcgg	cggctgacac	cagaacttgc	6840
153	ccacccctta	ttcgccggtg	qttttaagac	cggtaaaaga	tttctttagc	agaaaacact	6900
154	cggattaacc	ttgtcgtgct	ttcgcacatt	gcaaagtttt	ctcgcctgct	gcgccccata	6960
155	agacctagat	ccgtgtctca	gtacccatct	ccgggcctct	cctctcagag	cccgtatctg	7020
156	ttatageett	ggtgggccat	tacctcacca	acaagctgat	agaccgcagt	cccatcctac	7080
157	ggcgataaat	catttgggcc	acaaaccatt	ccaggcatag	tggcctatcg	gatattattc	7140
158	tcagtttccc	gaggttatcc	ccqtccatag	gttagattga	ctacgtgtta	ctgagccgtc	7200
159	taccttatat	tgctacaatg	actcgcatgg	cttagtatca	atccgatagc	agtcaggtcc	7260
160	ggcaggatca	accggattca	taattqqatt	atttttttt	tgttaagtac	gcttgtactt	7320
161	ttagaattaa	acagaatgca	cataatcttc	acatctcaga	tatgaccctt	cgatcatacc	7380
162	ctcattctgt	gtgcgtaact	ggaggccagc	gaatcacaat	atggtacaat	accatgcatt	7440
163	catcgcaage	gccgctcttg	catcacatac	gateggateg	gcccgtccat	gggcatataa	7500
164	accategeeg	atttccgccc	ccaacaaccc	cgatcagggg	ccggatctgc	ctgtatgatg	7560
165	gcgatccgcc	ctgattaaat	tatgggggga	acaacctact	gccgtggatc	tggaacgcga	7620
166	gtacaggga	aagaccggcg	actcaaccca	gatetttgee	aggtcgaaaa	agtaccacgt	7680
167	caacaaaatc	agccacaaca	taaggttcta	cgagccgtat	ccgtttgtga	caaggtccgc	7740
168	gaggggggaa	cacctcgtcg	acatagacag	gaacaagtat	gtagactact	ggatggggca	7800
169	ctagageeta	atactggggc	acgcgccggc	gccagtcagg	toggoagtag	aggggcagct	7860
170	teaccacaac	tggatccacg	ggaccgtcaa	cgagcagacg	atgaatctct	cggagataat	7920
171	acacaacaca	gtaagcgtgg	cagaaaagac	aaggtacgtc	acgtcgggga	cggaggccgt	7980
172	catgtatgcg	gcaaggctgg	cacacacaca	tacgggcaga	aaaataatag	caaaggcgga	8040
172	cacacactaa	cacgggtacg	catcagaact	gctcaagtcg	gtcaactggc	cgtatgatgt	8100
174	accedagage	ggggggctcg	togacgaaga	gcactctata	tccattccgt	acaacgatct	8160
175	tgaaggttcc	ctggatgttc	ttaggcgcgc	aggcgacgac	ttggcatgcg	tgataatcga	8220
176	accactacta	ggcggcggcg	gctgcatacc	ggcggatgag	gactatctgc	gcggcataca	8280
177	gaagtttata	cattcaaggg	acacactact	tatectegae	gagatagtga	cagggttccg	8340
178	ggugeeegeg	aactacacat	atgctgcagc	agggctggac	cccgatatag	tggcgctcgg	8400
179	caadatadto	gggggcggat	tececatage	ggtgatatgc	ggcaaggacg	aggtgatgga	8460
180	aatotooaac	actatatcgc	atgcaaagtc	cgacaggggg	tacatcggcg	geggeacatt	8520
181	ototocaac	cccgccacga	tgacagcggg	cacaacaaca	ctcggggagc	tcaaaaagag	8580
182	aaaaaaaaaa	atatacccga	ggataaactc	catagagaac	gacgcaaggg	acaagetete	8640
	aaagggcaca	gggaacaggg	tatccataac	Cadayadada	tcgctgttca	tgactcactt	8700
183	tattassast	gggaacaggg	gggtctcaaa	tactacagat.	gcggcagcct	gcgatgttga	8760
184	retretage	aggtagaga	tagacataat	Cacccadac	ggcatattct	ttctgccggg	8820
185	gergergeae	aggiactator	caacacacta	aaaggccgac	ctcaagacca	tgtattccgc	8880
186	caayetgggg	tttaaaaaaa	acatataaaa	tataggeegae	gaggaaactt	tgattatacg	8940
187	accagagege	accaracas	atratactor	toddcaadad	casccccacc	gagctggtgc	9000
188	ggegtgetge	act act at	acyacacicc	anttranger	ggcaataggc	ctgtacggga	9060
189	gccaggcgga	agaggagge	agcaagaacc	acat cotace	caaaaaaaaa	ctggcccaga	9120
190	aaatcctcaa	ggacgacccg	cayaacayyy	catacttta2	ccaactacta	gagettgaga	9180
191	acagggcaaa	aaaytactct	gatycgatca	cycyclicga	coggotgett	gagcttgaca	2200

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

192	acaaggacgc	gcccgcgtac	aacaacaagg	ccatagccca	ggccgagctc	ggagacacgg	9240
193	catccgcgct	ggaaaactac	ggcagggcca	tcgaggccga	cccgcggtac	gcgccggcgc	9300
194	gcttcaacag	ggccgtgctg	ctcgacaggc	tgggcgagca	tgaggaggcg	ctgccggacc	9360
195	tcgacagggc	agccgagctg	gaccgacgca	agccgaaccc	gaggttctac	aaggggatag	9420
196	tgctcggcaa	gatgggcagg	cacgaagagg	cgctggcctg	cttcaagggc	gtgtgcaaga	9480
197	ggcatcccgg	ccacgccgac	tcacagttcc	acgtggggat	agagettace	gagcttggca	9540
198	ggcacgccga	ggccctcggg	gagcttgcat	cactgcccgc	ggagcaccgc	gagaacgcca	9600
199	atgtattgta	tgccagggcg	cgcagcctct	cgggccttgg	cagggaggac	gaatccatag	9660
200	cgcacctgca	aaaggcggcc	aaaaaagatt	ccaagacgat	aaaaaagtgg	gcccgcgcag	9720
201	aaaaggcctt	tgacggaata	cgggacgatc	ccggttcaaa	aagatagccg	gctagaggat	9780
202	cttttttctt	gccgcgtcaa	tccgcatcat	gcggaccttt	tttttgggcc	ccacaagtcg	9840
203	cgattcatag	actggtacat	agaccacctc	caccgccttt	gcggcaaact	cctcccgcag	9900
204	gtcgcgcatg	ccgtcaggcg	ggggcccgcg	cagcttctct	tttagttttg	agagcgcctc	9960
205	ttctgtctcc	acctcggggc	tccgcacatt	ctctgacgca	tcgagtatcc	tccgcgggta	10020
206	cggctccacc	gcgccgggcc	ccgtcttgta	gggaaagtcc	gtctcgccgc	cgtgccggtc	10080
207	aaggcacatc	atcccttctg	attccgcaaa	gacatgctct	tctagctcga	ggtcgaccct	10140
208	gttcttgccg	agccctgccg	agagcgtctt	gtgtatgcgc	gacttggacc	ttatgggaaa	10200
209	gacgccgtcg	cctagcacca	cctcgatcac	gttctggtcc	accttgatcg	ggtgaaccgc	10260
210	ctttctgaaa	aaatccgcag	agtacctggc	ggagacccgg	atcagcgcct	cgtggaccag	10320
211	ctttacagaa	tgcacatgga	cgtcttcttt	ccgcggggcc	ctcataaggg	ccctaaaggc	10380
212	acceptette	tttgcctcta	tcatggcccg	agccgactcc	tcagtcatgg	cgttccgcag	10440
213	gaccgccgtc	ctqqtctttc	cagtcatccc	ctgccgcacc	ccgcataagg	catactatac	10500
214	aacqcaaqqc	aaqqtaataa	tagcctgccg	tctgtaacgg	ccgtatgagg	tcggagggca	10560
215	ggcccggata	catcgaaaag	ttcctaaaga	gggcggacaa	ggcgatagac	aatgcagtcg	10620
216	agcagggggt	caaqaqqqca	gacgagatac	tagatgacgc	agtcgagctc	ggcaagatca	10680
217	ccatagacaa	ggcgcaaaaa	agaagcgatg	tgctgctcaa	gcaggccgag	cgggagagca	10740
218	agcggctcaa	gtcaaggggc	gccaaaaagc	tcgaaaaggg	cataggggcg	gcaaaaaaga	10800
219	taacaaccaa	caagggcgac	gcgctagaga	ccctggcaaa	gctcggcgag	ctgagaaagg	10860
220	cggggatcat	aacggagaag	gagtttcgcg	ccaaqaaaaa	gaagcttctc	gcggagatct	10920
221	gacttgaage	coctagacta	tacccgggac	ggctcgataa	aggaggtcac	aaagaggtgg	10980
222	tttataggca	caccatccct	tgtcgacctt	gcaggcgagc	toggcatato	tgagagcaac	11040
223	atattccacq	tgacatttcc	cgacggcgca	aaqaccaccc	tgcatacgca	cgagggcggg	11100
224	cagetgetea	tagtgacctc	gggcaccggc	agcatgtcaa	tatttgaaaa	gaccggcgga	11160
225	aacaaaacaa	aatttgcaat	aaaaqaqaca	gacaggatac	cgctaaagca	gggcagcatc	11220
226	cagtacatac	ctaccaacat	gctcacgtgc	acqqcqcaac	agacggcacc	accetgteee	11280
227	atatagcggt	aaactacccg	tcgccatcgg	gaaaggagcc	gtatacatta	tggtatgaat	11340
228	ccgactttgc	cagccgggtc	accogcotoc	totaaattat	attatttgag	cctctccagt	11400
229	atcgacagge	ttacaaggtt	ggtcatcgtt	atccccttqc	ggatcacttc	ccttctagtc	11460
230	ttttcgcagt	acttgttgac	actctaataa	ctcttttcqc	tggatacctc	gagcaccaca	11520
231	atottcccgc	tgtacgggaa	ggcaaacttt	ggcggaatcc	togagoatao	atccatgcgc	11580
232	cctattcccq	ccctgcctat	cttctcctt	cttgccacta	ccgacqccac	ctcgcgtatg	11640
233	tectecteg	aatctccgta	ttccagatag	tacatggata	catageteat	cccgggggat	11700
234	tecettegg	atateteete	gtccatgctg	aataaatagg	aagggccgcc	cccgcggccc	11760
235	tocaccacct	ttatcatttt	gaaaccattt	ttgaaccttg	ccagcaggta	gtggctggcc	11820
235	tactactas	agtacggctt	gettttagae	gagaaagtgg	cggtcacaaa	gtacatctcg	11880
237	cotacettee	togacoagoet	ccattcctcc	ttgageteta	tegaggeatg	gtcgtgcgtg	11940
237	tataaaataa	aggestates	CCCCaaaasa	acctccatct	tggacatgca	tgggatccgc	12000
238	daaggegtge	aatatotaot	tecataceae	cttagagaga	adadacccca	cctgtggccg	12060
239	ggaaccggtt	aggegtaegt	ggatccatge	gatagttatt	taaaactagg	atgccgatca	12120
240	goodaygggd	aggegegege	ggaccoacgo	5			



VERIFICATION SUMMARY

PATENT APPLICATION: US/10/027,806

DATE: 01/18/2002 TIME: 09:17:39

Input Set : N:\Crf3\RULE60\10027806.raw
Output Set: N:\CRF3\01182002\J027806.raw

L:9 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD

L:7288 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:120